PERGAMON Computers and Mathematics with Applications 44 (2002) 407–411

An International Journal
computers & mathematics with applications

www.elsevier.com/locate/camwa

Relative Errors Versus Residuals of Approximate Solutions of Weighted Least Squares Problems in Hilbert Space

JIU DING
Department of Mathematics
The University of Southern Mississippi
Hattiesburg, MS 39406-5045, U.S.A.

YIMIN WEI
Department of Mathematics
Fudan University
Shanghai, 200433, P.R. China

(Received February 2001; revised and accepted August 2001)

Abstract—We give some lower and upper bounds in terms of residuals for the perturbation of general consistent systems of linear operator equations and weighted least squares problems in Hilbert space.

Keywords—Weighted least squares, Weighted generalized inverse.

1. INTRODUCTION

Let X and Y be Hilbert spaces, let $A : X \to Y$ be a bounded linear operator with closed range, and let $b \in Y$ be a fixed vector.

We consider the nonhomogeneous system of linear operator equations in Hilbert space,

$$Ax = b, \tag{1}$$

and its perturbed system

$$(A + E)y = b + e. \tag{2}$$

When A in (1) is invertible, a well-known perturbation result (see, e.g., inequalities (2.3.3) and (2.3.11) in [1]) says that if y approximates the exact solution $x = A^{-1}b$ of (1), then

$$\frac{1}{\kappa}\frac{\|r_y\|}{\|b\|} \le \frac{\|y - x\|}{\|x\|} \le \kappa\frac{\|r_y\|}{\|b\|}, \tag{3}$$

We thank the two referees for good comments and suggestions. Research of the second author was supported by the National Natural Science Foundation of China (19901006), and partial work was finished when the author visited Harvard University and was supported by the China Scholarship Council.

0898-1221/02/$ - see front matter Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX
PII: S0898-1221(02)00158-X

where $r_y = Ay - b$ is the *residual* associated with y and $\kappa = \|A\|\|A^{-1}\|$ is the *condition number* of A with $\|\ \|$ standing for the operator norm corresponding to any vector norm on X and Y (not necessarily the one induced by the inner product).

Recently, the inequalities in (3) have been generalized in [2] for general matrices, with the help of generalized inverse technique. In this paper, we further extend the known results to weighted least squares problems in Hilbert space, using the concept of weighted generalized inverses of operators. Since our main results, Theorem 2.1 and Theorem 3.1, work for general cases in which systems (1) and (10) usually have infinitely many solutions, unlike the classic result (3) in which the relative error is defined as the absolute error divided by the norm of the exact solution, we introduce the generalized concept of relative error as the ratio of the absolute error to the distance of an exact solution to the null space of A. Lower and upper bounds of such relative errors will be given by formulas (4) and (12) in the following for consistent systems and least squares problems, respectively.

In the next section, we consider the case when (1) is consistent, and the most general case of weighted least squares problems in Hilbert space will be dealt with in Section 3.

2. THE CASE WHEN $b \in R(A)$ IN HILBERT SPACE

Let $N(A)$ and $R(A)$ be the null space and the range of A, respectively. We need the concept of the *weighted generalized inverse* $A^{\dagger}_{MN}$ of A, which is defined as the unique linear operator $A^{\dagger}_{MN}$ satisfying

$$AA^{\dagger}_{MN}A = A, \qquad A^{\dagger}_{MN}AA^{\dagger}_{MN} = A^{\dagger}_{MN},$$
$$\left(NA^{\dagger}_{MN}A\right)^* = NA^{\dagger}_{MN}A, \qquad \left(MAA^{\dagger}_{MN}\right)^* = MAA^{\dagger}_{MN},$$

where $M : Y \to Y$ and $N : X \to X$ are positive definite linear operators. In particular, when $X = R^n$, $Y = R^m$, and $M = I_m$, $N = I_n$, the weighted generalized inverse is reduced to the regular generalized inverse of matrices. When the norms for X and Y are induced from their respective inner products, e.g., the Euclidean norm $\|\cdot\|_2 \equiv (\cdot,\cdot)^{1/2}$, then $A^{\dagger}_{MN}$ is called the weighted Moore-Penrose generalized inverse of A. See [3–5] for more details on $A^{\dagger}_{MN}$.

In this section, we assume that a norm $\|\ \|$ is given on X and Y and the operator norm is the induced one. Again, let $\kappa = \|A\|\|A^{\dagger}_{MN}\|$ be the condition number of A. Let $d(x, N(A)) = \inf\{\|x - z\| : z \in N(A)\}$ be the distance of x to $N(A)$ and let $r_y = Ay - b$ be the residual associated with y.

THEOREM 2.1. *Suppose $b \in R(A)$. Then for any $y \in X$, there is a solution x to (1) such that*

$$\frac{1}{\kappa}\frac{\|r_y\|}{\|b\|} \le \frac{\|y - x\|}{\left\|A^{\dagger}_{MN}b\right\|} \le \frac{\|y - x\|}{d(x, N(A))} \le \kappa\frac{\|r_y\|}{\|b\|}. \tag{4}$$

PROOF. Let $x = A^{\dagger}_{MN}b + (I - A^{\dagger}_{MN}A)y$ be the projection of y onto the solution set of (1) along $N(A)$. Then

$$\left\|A^{\dagger}_{MN}b\right\| = \left\|x - \left(I - A^{\dagger}_{MN}A\right)y\right\| \ge d(x, N(A)),$$

so the middle inequality of (4) is valid. Since $Ax = b$, we have

$$A(y - x) = Ay - b. \tag{5}$$

On the other hand, by the construction of x,

$$y - x = A^{\dagger}_{MN}(Ay - b). \tag{6}$$

Let $z \in N(A)$ be arbitrary. Then (6) gives that

$$\frac{\|y-x\|}{\|x-z\|} \le \frac{\left\|A_{MN}^{\dagger}\right\| \|Ay-b\|}{\|x-z\|} = \frac{\|A\| \left\|A_{MN}^{\dagger}\right\| \|r_y\|}{\|A\|\|x-z\|}$$
$$\le \frac{\kappa\|r_y\|}{\|Ax\|} = \kappa \frac{\|r_y\|}{\|b\|},$$

which gives the right inequality of (4). Now (5) implies that

$$\frac{\|r_y\|}{\|b\|} = \frac{\|Ay-b\|}{\|b\|} \le \frac{\|A\|\|y-x\|}{\|b\|}$$
$$\le \frac{\left\|A_{MN}^{\dagger}\right\| \|A\|\|y-x\|}{\left\|A_{MN}^{\dagger}\right\| \|b\|} \le \kappa \frac{\|y-x\|}{\left\|A_{MN}^{\dagger}b\right\|}.$$

This gives the left inequality of (4). ∎

REMARK 2.1. When the norms $\| \ \|$ are given via the inner products, $\|y-x\|$ is exactly the minimal distance of y to the affine set of all solutions to (1), and in this case, (4) gives lower and upper bounds of this distance with respect to the distance of the solution x to the null space of A.

Since A^*A is invertible implies that $N(A) = \{0\}$ and $x = A_{MN}^{\dagger}b$, we immediately have the following corollary.

COROLLARY 2.1. *If in addition, A^*A is invertible, then*

$$\frac{1}{\kappa}\frac{\|r_y\|}{\|b\|} \le \frac{\left\|y - A_{MN}^{\dagger}b\right\|}{\left\|A_{MN}^{\dagger}b\right\|} \le \kappa \frac{\|r_y\|}{\|b\|}. \tag{7}$$

In particular, if A is nonsingular, then (4) is reduced to (3).

The residual r_y depends on E and e if y turns out a solution of (2). In this case, subtracting (1) from (2) gives that

$$r_y = A(y-x) = e - Ey,$$

from which it follows that

$$\|e - Ey\| = \|r_y\| \le \|e\| + \|E\|\|y\|. \tag{8}$$

Therefore, we have the following corollary.

COROLLARY 2.2. *If in addition, $b + e \in R(A+E)$, then for any solution y to (2), there is a solution x to (1) such that*

$$\frac{1}{\kappa}\frac{\|e-Ey\|}{\|b\|} \le \frac{\|y-x\|}{\left\|A_{MN}^{\dagger}b\right\|} \le \frac{\|y-x\|}{d(x,N(A))} \le \kappa \frac{\|e\| + \|E\|\|y\|}{\|b\|}. \tag{9}$$

3. WEIGHTED LEAST SQUARES PROBLEMS IN HILBERT SPACE

Now we consider the weighted least squares problem

$$\|Ax - b\|_M = \min_{z \in X} \|Az - b\|_M \tag{10}$$

and its perturbed form

$$\|(A+E)y-(b+e)\|_M = \min_{z\in X} \|(A+E)z-(b+e)\|_M. \tag{11}$$

Here, we assume that the weighted norms on Y and X are determined by inner products, i.e., $\|y\|_M^2 = (My,y)$ for $y \in Y$, and $\|x\|_N^2 = (Nx,x)$ for $x \in X$. It is well known [3] that $A_{MN}^\dagger b$ is a solution of (10). The condition number with respect to the weighted norms is defined as $\kappa_{MN} = \|A\|_{MN}\|A_{MN}^\dagger\|_{NM}$. Let R be a positive definite linear operator on a Hilbert space Z and S be a positive definite linear operator on a Hilbert space W. Then for a given bounded linear operator T from Z to W, we define an operator $T^\#$ from W to Z by $T^\# = R^{-1}T^*S$.

THEOREM 3.1. *For any $y \in X$, there is a solution x to (10) such that*

$$\frac{1}{\kappa_{MN}} \frac{\|r_y\|_M}{\left\|AA_{MN}^\dagger b\right\|_M} \le \frac{\|y-x\|_N}{\left\|A_{MN}^\dagger b\right\|_N} = \frac{\|y-x\|_N}{d(x,N(A))} \le \kappa_{MN} \frac{\|r_y\|_M}{\left\|AA_{MN}^\dagger b\right\|_M}, \tag{12}$$

where $r_y = Ay - AA_{MN}^\dagger b$ and $d(x,N(A))$ is defined with respect to the norm $\|\cdot\|_N$.

PROOF. Let $x = A_{MN}^\dagger b + (I - A_{MN}^\dagger A)y$ be the projection of y onto the solution set of (10) along $N(A)$. Then the middle equality of (12) is obvious. Now, as in the previous section,

$$y - x = A_{MN}^\dagger(Ay-b) = A_{MN}^\dagger\left(Ay - AA_{MN}^\dagger b\right), \tag{13}$$

from which it follows that

$$A(y-x) = AA_{MN}^\dagger\left(Ay - AA_{MN}^\dagger b\right) = Ay - AA_{MN}^\dagger b. \tag{14}$$

The proof from here on is exactly the same as that for Theorem 2.1 except that b is replaced with $AA_{MN}^\dagger b$, so it will be omitted. ∎

Now we estimate $\|r_y\|_M$ if y solves (11). Let

$$Ax - b = r, \qquad (A+E)y-(b+e) = r', \tag{15}$$

and subtract the first equality from the second one. We have

$$r_y = A(y-x) = e - Ey + r' - r. \tag{16}$$

Since $A_{MN}^\dagger r = 0$ and $(A+E)_{MN}^\dagger r' = 0$, (13) and (16) imply that

$$\begin{aligned} y - x &= A_{MN}^\dagger r_y = A_{MN}^\dagger(e - Ey + r' - r) = A_{MN}^\dagger(e - Ey + r') \\ &= A_{MN}^\dagger(e-Ey) + \left[A_{MN}^\dagger - (A+E)_{MN}^\dagger\right] r'. \end{aligned} \tag{17}$$

By the decomposition formula (Theorem 8.5 of [4]) that

$$\begin{aligned} A_{MN}^\dagger - (A+E)_{MN}^\dagger = A_{MN}^\dagger E(A+E)_{MN}^\dagger - A_{MN}^\dagger\left(A_{MN}^\dagger\right)^\# E^\# \left[I - (A+E)(A+E)_{MN}^\dagger\right] \\ - \left(I - A_{MN}^\dagger A\right) E^\# \left[(A+E)_{MN}^\dagger\right]^\# (A+E)_{MN}^\dagger, \end{aligned}$$

where $A^\# = N^{-1}A^*M$, we obtain

$$\left[A_{MN}^\dagger - (A+E)_{MN}^\dagger\right] r' = -A_{MN}^\dagger\left(A_{MN}^\dagger\right)^\# E^\# r', \tag{18}$$

from which (17) can be written as

$$y - x = A_{MN}^{\dagger}(e - Ey) - A_{MN}^{\dagger}\left(A_{MN}^{\dagger}\right)^{\#} E^{\#} r'. \tag{19}$$

It follows from (14) and (19) that

$$\begin{aligned} r_y = A(y-x) &= AA_{MN}^{\dagger}(e - Ey) - AA_{MN}^{\dagger}\left(A_{MN}^{\dagger}\right)^{\#} E^{\#} r' \\ &= AA_{MN}^{\dagger}(e - Ey) - \left(EA_{MN}^{\dagger}\right)^{\#} r'. \end{aligned} \tag{20}$$

On the other hand, from the definition of the weighted norm,

$$\begin{aligned} \left\|\left(EA_{MN}^{\dagger}\right)^{\#}\right\|_{MM} &= \left\|M^{1/2}\left(EA_{MN}^{\dagger}\right)^{\#} M^{-1/2}\right\|_2 = \left\|M^{1/2}M^{-1}\left(EA_{MN}^{\dagger}\right)^{*} MM^{-1/2}\right\|_2 \\ &= \left\|M^{-1/2}\left(EA_{MN}^{\dagger}\right)^{*} M^{1/2}\right\|_2 = \left\|M^{1/2}EA_{MN}^{\dagger}M^{-1/2}\right\|_2 = \left\|EA_{MN}^{\dagger}\right\|_{MM}. \end{aligned}$$

Therefore,

$$\begin{aligned} \|r_y\|_M &\le \|e\|_M + \|E\|_{MN}\|y\|_N + \left\|\left(EA_{MM}^{\dagger}\right)^{\#}\right\|_{MM} \|r'\|_M \\ &= \|e\|_M + \|E\|_{MN}\|y\|_N + \left\|EA_{MM}^{\dagger}\right\|_{MM} \|r'\|_M. \end{aligned}$$

Thus, we have the following corollary.

COROLLARY 3.1. *For any solution y to (11), there is a solution x to (10) such that*

$$\begin{aligned} &\frac{1}{\kappa_{MN}} \frac{\left\|AA_{MN}^{\dagger}(e - Ey) - \left(EA_{MN}^{\dagger}\right)^{\#} r'\right\|_M}{\left\|AA_{MN}^{\dagger}b\right\|_M} \le \frac{\|y - x\|_N}{\left\|A_{NM}^{\dagger}b\right\|_N} \\ &\le \frac{\|y - x\|_N}{d(x, N(A))} \le \kappa_{MN} \frac{\|e\|_M + \|E\|_{MN}\|y\|_N + \left\|EA_{MN}^{\dagger}\right\|_{MM} \|r'\|_M}{\left\|AA_{MN}^{\dagger}b\right\|_M}. \end{aligned} \tag{21}$$

REFERENCES

1. D.S. Watkins, *Fundamentals of Matrix Computations*, John Wiley & Sons, (1991).
2. J. Ding, Relative errors versus residuals of approximate solutions to linear algebraic equations and least squares problems, *Missouri J. Math. Sci.* **13** (1), 47–52, (2001).
3. S. Campbell and C. Meyer, *Generalized Inverses of Linear Transformations*, Pitman, (1979).
4. C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, NJ, (1974).
5. Y. Wei and H. Wu, Expression for the perturbation of the weighted Moore-Penrose inverse, *Computers Math. Applic.* **39** (5/6), 13–18, (2000).